

Hyperconvexity and Tight Span Theory for Diversities

David Bryant^{a,*}, Paul F. Tupper^b

^a*Dept. of Mathematics and Statistics, University of Otago. PO Box 56 Dunedin 9054, New Zealand. Ph (64)34797889. Fax (64)34798427*

^b*Dept. of Mathematics, Simon Fraser University. 8888 University Drive, Burnaby, British Columbia V5A 1S6, Canada. Ph (778)7828636. Fax (778)7824947*

Abstract

The tight span, or injective envelope, is an elegant and useful construction that takes a metric space and returns the smallest hyperconvex space into which it can be embedded. The concept has stimulated a large body of theory and has applications to metric classification and data visualisation. Here we introduce a generalisation of metrics, called diversities, and demonstrate that the rich theory associated to metric tight spans and hyperconvexity extends to a seemingly richer theory of diversity tight spans and hyperconvexity.

Keywords: Tight span; Injective hull; Hyperconvex; Diversity; Metric geometry;

1. Introduction

Hyperconvex metric spaces were defined by Aronszajn and Panitchpakdi in [1] as part of a program to generalise the Hahn-Banach theorem to more general metric spaces (reviewed in [2], and below). Isbell [3] and Dress [4] showed that every metric space could be embedded into a minimum hyperconvex space, called the *tight span* or *injective envelope*. Our aim is to show that the notion of hyperconvexity, the tight span, and much of the related theory can be extended beyond metrics to a class of multi-way metrics which we call diversities.

*Corresponding author

Email addresses: david.bryant@otago.ac.nz (David Bryant), pft3@math.sfu.ca (Paul F. Tupper)

Recall that a *metric space* is a pair (X, d) where X is a set and d is a function from $X \times X$ to \mathfrak{R} satisfying

(M1) $d(a, b) \geq 0$ and $d(a, b) = 0$ if and only if $a = b$.

(M2) $d(a, c) \leq d(a, b) + d(b, c)$.

We define a *diversity* to be a pair (X, δ) where X is a set and δ is a function from the finite subsets of X to \mathfrak{R} satisfying

(D1) $\delta(A) \geq 0$, and $\delta(A) = 0$ if and only if $|A| \leq 1$.

(D2) If $B \neq \emptyset$ then $\delta(A \cup C) \leq \delta(A \cup B) + \delta(B \cup C)$.

We prove below that these axioms imply monotonicity:

(D3) If $A \subseteq B$ then $\delta(A) \leq \delta(B)$.

We will show that tight span theory adapts elegantly from metric spaces to diversities. The tight span of a metric space (X, d) is formed from the set of point-wise minimal functions $f : X \rightarrow \mathfrak{R}$ such that $f(a_1) + f(a_2) \geq d(a_1, a_2)$ for all $a_1, a_2 \in X$. Letting $\mathcal{P}_{\text{fin}}(X)$ denote the finite subsets of X , the tight span of a diversity (X, δ) is formed from the set of point-wise minimal functions $f : \mathcal{P}_{\text{fin}}(X) \rightarrow \mathfrak{R}$ such that

$$f(A_1) + f(A_2) + \cdots + f(A_k) \geq \delta(A_1 \cup A_2 \cup \cdots \cup A_k)$$

for all finite collections $\{A_1, A_2, \dots, A_k\} \subseteq \mathcal{P}_{\text{fin}}(X)$. We can embed a metric space in its tight span (with the appropriate metric on the tight span); we can embed a diversity in its tight span (with the appropriate diversity on the tight span). Both constructions have characterisations in terms of injective hulls, and both possess a rich mathematical structure.

The motivation for exploring tight spans of diversities was the success of the metric tight span as a tool for classifying and visualising finite metrics, following the influential paper of Dress [4]. The construction provided the theoretical framework for split decomposition [5] and Neighbor-Net [6], both implemented in the SplitsTree package [7] and widely used for visualising phylogenetic data.

By looking at diversities, rather than metrics or distances, our hope is to incorporate more information into the analysis and thereby improve inference [8].

Dress et al. [9] coined the term *T-theory* for the field of discrete mathematics devoted to the combinatorics of the tight span and related constructions. Sturmfels [10] highlighted T-theory as one area where problems from biology have led to substantial new ideas in mathematics. Contributions to T-theory include profound results on optimal graph realisations of metrics [4, 11, 12]; intriguing connections between the Buneman graph, the tight span and related constructions [9, 12–16]; links with tropical geometry and hyperdeterminants [17, 18]; classification of finite metrics [4, 19]; and properties of the tight span for special classes of metrics [20, 21]. Hirai [22] describes an elegant geometric formulation of the tight span. We believe that there will be diversity analogues for many of these metric space results.

Our use of the term *diversity* comes from the appearance of a special case of our definition in work on phylogenetic and ecological diversity [8, 23–25]. However diversities crop up in a broad range of contexts, for example:

1. *Diameter Diversity.* Let (X, d) be a metric space. For all $A \in \mathcal{P}_{\text{fin}}(X)$ let

$$\delta(A) = \max_{a, b \in A} d(a, b) = \text{diam}(A).$$

Then (X, δ) is a diversity.

2. *L_1 diversity.* For all finite $A \subseteq \mathbb{R}^n$ define

$$\delta(A) = \sum_i \max_{a, b} \{|a_i - b_i| : a, b \in A\}.$$

Then (\mathbb{R}^n, δ) is a diversity.

3. *Phylogenetic Diversity.* Let T be a phylogenetic tree with taxon set X . For each finite $A \subseteq X$, let $\delta(A)$ denote the length of the smallest subtree of T connecting taxa in A . Then (X, δ) is a *phylogenetic diversity*.
4. *Length of the Steiner Tree.* Let (X, d) be a metric space. For each finite $A \subseteq X$ let $\delta(A)$ denote the minimum length of a Steiner tree connecting elements in A . Then (X, δ) is a diversity.

5. *Truncated diversity.* Let (X, δ) be a diversity. For all $A \in \mathcal{P}_{\text{fin}}(X)$ define

$$\delta^{(k)}(A) = \max\{\delta(B) : |B| \leq k, B \subseteq A\}.$$

For each $k \geq 2$, $(X, \delta^{(k)})$ is a diversity. Note that these diversities can be encoded using $O(|X|^k)$ values, an important consideration for the design of efficient algorithms.

The generalisation of metrics to more than two arguments has a long history. There is an extensive literature on 2-metrics (metrics taking three points as arguments); see [26]. Generalised metrics defined on n -tuples for arbitrary n go back at least to Menger [27], who took the volume of an n -simplex in Euclidean space as the prototype. Recently various researchers have continued the study of such generalised metrics defined on n -tuples; see [28–30] for examples. However, as of yet, a satisfactory theory of tight spans has not been developed for these generalisations.

Dress and Terhalle [31] developed tight span theory for *valuated matroids*, which can be viewed as an n -dimensional version of a restricted class of metrics. They demonstrated intriguing links with algebraic building theory. One significant difference is that, for diversities, the tight span consists of functions on $\mathcal{P}_{\text{fin}}(X)$ rather than on X , as is the case for valuated matroids.

We note that our results differ from all of this earlier work because, for a diversity (X, δ) , the function δ is defined on arbitrary finite subsets of X rather than tuples of a fixed length.

The structure of this paper is as follows: In Section 2 we develop the basic theory of tight spans on diversities, defining the appropriate diversity for a tight span and showing that every diversity embeds into its tight span. In Section 3 we characterise diversities that are isomorphic to their tight spans. Here, isomorphism is defined in analogy to isometry for metric spaces. These are the *hyperconvex* diversities, a direct analogue of metric hyperconvexity. We prove that diversity tight spans, like metric tight spans, are injective, and are formally the injective envelope in the category of diversities. In Section 4 we explore in more detail the direct links between diversity tight spans and metric tight spans.

We show when the diversity equals the diameter diversity (as defined above) the diversity tight span is isomorphic to the diameter of the metric tight span. In Section 5 we study the tight span of a phylogenetic diversity, and prove that taking the tight span of a phylogenetic diversity recovers the underlying tree in the same way that taking the tight span of an additive metric recovers its underlying tree. This theory is developed for *real trees*. Finally, in Section 6 we examine applications of the theory to the classical Steiner Tree problem. Dress and Krüger [32] defined an *abstract* Steiner tree where the internal nodes did not have to sit in the given metric space. They proved that these abstract Steiner trees can be embedded in the tight span. We extend their results to Steiner trees based on diversities, thereby obtaining tight bounds for the classical Steiner tree problem.

2. The tight span of a diversity

We begin by establishing some basic properties of diversities. Recall that $\mathcal{P}_{\text{fin}}(X)$ denotes all the finite subsets of the set X and that a diversity is the pair (X, δ) where X is a set and the function $\delta: \mathcal{P}_{\text{fin}}(X) \rightarrow \mathfrak{R}$ satisfies axioms (D1) and (D2).

Proposition 2.1. *Let (X, δ) be a diversity.*

1. *If $d: X \times X \rightarrow \mathfrak{R}$ is defined as $d(x, y) = \delta(\{x, y\})$ then (X, d) is a metric space. We say that (X, d) is the induced metric of (X, δ) .*
2. *(D3) holds, that is, for $A, B \in \mathcal{P}_{\text{fin}}(X)$, if $A \subseteq B$ then $\delta(A) \leq \delta(B)$.*
3. *For $A, B \in \mathcal{P}_{\text{fin}}(X)$ if $A \cap B \neq \emptyset$ then $\delta(A \cup B) \leq \delta(A) + \delta(B)$.*

Proof.

1. That $d(x, y) = 0$ if and only if $x = y$ follows from (D1). Symmetry is clear. To obtain the triangle inequality, for any $x, y, z \in X$,

$$d(x, z) = \delta(\{x, z\}) \leq \delta(\{x, y\}) + \delta(\{y, z\}) = d(x, y) + d(y, z),$$

using (D2).

2. First note for any $a \in A$ and $b \in X$ that by (D2) with C empty

$$\delta(A) \leq \delta(A \cup \{b\}) + \delta(\{b\}) = \delta(A \cup \{b\}).$$

The more general result follows by induction.

3. Using (D2)

$$\delta(A \cup B) \leq \delta(A \cup (A \cap B)) + \delta(B \cup (A \cap B)) = \delta(A) + \delta(B).$$

□

We now state the diversity analogue for the metric tight span.

Definition 2.2. Let (X, δ) be a diversity. Let P_X denote the set of all functions $f: \mathcal{P}_{\text{fin}}(X) \rightarrow \mathfrak{R}$ satisfying $f(\emptyset) = 0$ and

$$\sum_{A \in \mathcal{A}} f(A) \geq \delta\left(\bigcup_{A \in \mathcal{A}} A\right) \quad (2.1)$$

for all finite $\mathcal{A} \subseteq \mathcal{P}_{\text{fin}}(X)$. Write $f \preceq g$ if $f(A) \leq g(A)$ for all finite $A \subseteq X$. The tight span of (X, δ) is the set T_X of functions in P_X that are minimal under \preceq .

Example. Any diversity δ on $X = \{1, 2, 3\}$ is determined by the four values

$$d_{12} = \delta(\{1, 2\}), \quad d_{23} = \delta(\{2, 3\}), \quad d_{13} = \delta(\{1, 3\}), \quad d_{123} = \delta(\{1, 2, 3\}).$$

We write $f_i = f(\{i\})$, $f_{ij} = f(\{i, j\})$ and $f_{123} = f(\{1, 2, 3\})$ for $i, j \in X$.

Condition (2.1) then translates to the following set of inequalities:

$$\begin{aligned} f_i &\geq 0 \\ f_{ij} &\geq d_{ij} \\ f_i + f_j &\geq d_{ij} \\ f_{123} &\geq d_{123} \\ f_i + f_{jk} &\geq d_{123} \\ f_1 + f_2 + f_3 &\geq d_{123} \end{aligned} \quad (2.2)$$

for distinct $i, j, k \in X$. Note we have omitted inequalities like $f_{ij} + f_{jk} \geq d_{123}$ since these are implied by (2.2) and the triangle inequality (D2). The elements of T_X are the minimal f in P_X . Equivalently, T_X is the set of f that satisfy (2.2) and such that for each nonempty $A \subseteq X$, f_A appears in an inequality in (2.2) that is tight.

A straightforward but tedious analysis of the inequalities (which we omit) gives the following characterisation of T_X . Define the three ‘external’ vertices

$$\begin{aligned} v^{(1)} &= (0, d_{12}, d_{13}) \\ v^{(2)} &= (d_{12}, 0, d_{23}) \\ v^{(3)} &= (d_{13}, d_{23}, 0) \end{aligned}$$

and the four ‘internal’ vertices

$$\begin{aligned} u^{(0)} &= (d_{123} - d_{23}, d_{123} - d_{13}, d_{123} - d_{12}). \\ u^{(1)} &= u^{(0)} - (\beta, 0, 0) \\ u^{(2)} &= u^{(0)} - (0, \beta, 0) \\ u^{(3)} &= u^{(0)} - (0, 0, \beta), \end{aligned}$$

where $\beta = \max(2d_{123} - d_{12} - d_{23} - d_{13}, 0)$. Let C be the cell complex formed from the line segments $[u^{(1)}, v^{(1)}]$, $[u^{(2)}, v^{(2)}]$, $[u^{(3)}, v^{(3)}]$ and the solid tetrahedron with vertices $u^{(1)}, \dots, u^{(4)}$. A case-by-case analysis gives that $f \in T_X$ if and only if $(f_1, f_2, f_3) \in C$, $f_{23} = \max(d_{23}, d_{123} - f_1)$, $f_{13} = \max(d_{13}, d_{123} - f_2)$, $f_{12} = \max(d_{12}, d_{123} - f_3)$, $f_{123} = d_{123}$. If $\beta = 0$ then $u^{(0)}$ to $u^{(3)}$ coincide, and the tight span is one-dimensional and resembles the metric tight span for the induced metric, albeit sitting in a higher dimensional space (Figure 1a). When $\beta > 0$ the tight span resembles a tetrahedron with three spindles branching off, as in Figure 1b. \square

We now prove a characterisation of the diversity tight span which will be used extensively throughout the remainder of the paper (Theorem 2.3). An equivalent result holds for the metric tight span [4, Theorem 3(v)].

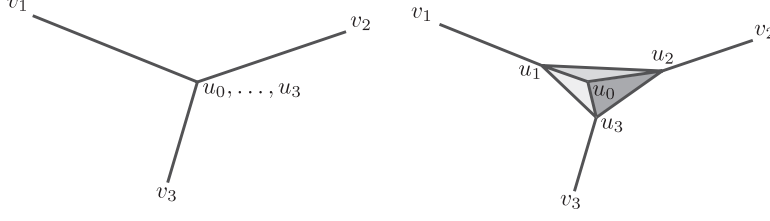


Figure 1: Two examples of the tight span on three points, with different values for $d(\{1, 2, 3\})$. On the left an example where $2d_{123} \leq d_{12} + d_{23} + d_{13}$, and the diversity tight span is one-dimensional and resembles the tight span of the induced metric. On the right a case with $2d_{123} > d_{12} + d_{23} + d_{13}$, where the diversity consists of a three-cell with three adjacent one-cells.

Theorem 2.3. *Let $f: \mathcal{P}_{\text{fin}}(X) \rightarrow \mathfrak{R}$ and suppose $f(\emptyset) = 0$. Then $f \in T_X$ if and only if for all finite $A \subseteq X$,*

$$f(A) = \sup_{\mathcal{B} \subseteq \mathcal{P}_{\text{fin}}(X)} \left\{ \delta\left(A \cup \bigcup_{B \in \mathcal{B}} B\right) - \sum_{B \in \mathcal{B}} f(B) : |\mathcal{B}| < \infty \right\}. \quad (2.3)$$

Proof.

Suppose that $f \in T_X$. As $f(\emptyset) = 0$, the supremum in (2.3) is well defined. For all finite $A \subseteq X$ and all finite $\mathcal{B} \subseteq \mathcal{P}_{\text{fin}}(X)$ we have

$$f(A) \geq \delta\left(A \cup \bigcup_{B \in \mathcal{B}} B\right) - \sum_{B \in \mathcal{B}} f(B),$$

giving the required lower bound on $f(A)$. Now suppose that for some finite A_0

$$f(A_0) > \sup_{\mathcal{B} \subseteq \mathcal{P}_{\text{fin}}(X)} \left\{ \delta\left(A_0 \cup \bigcup_{B \in \mathcal{B}} B\right) - \sum_{B \in \mathcal{B}} f(B) : |\mathcal{B}| < \infty \right\}. \quad (2.4)$$

Define a function $g: \mathcal{P}_{\text{fin}}(X) \rightarrow \mathfrak{R}_{\geq 0}$ by

$$g(A) = \begin{cases} f(A) & \text{if } A \neq A_0 \\ \sup_{\mathcal{B} \subseteq \mathcal{P}_{\text{fin}}(X)} \left\{ \delta\left(A_0 \cup \bigcup_{B \in \mathcal{B}} B\right) - \sum_{B \in \mathcal{B}} f(B) : |\mathcal{B}| < \infty \right\} & \text{if } A = A_0. \end{cases}$$

Clearly $g \neq f$ and $g \preceq f$. We show that g is in P_X . Let \mathcal{A} be a finite subset of $\mathcal{P}_{\text{fin}}(X)$. If $A_0 \notin \mathcal{A}$ then

$$\sum_{A \in \mathcal{A}} g(A) = \sum_{A \in \mathcal{A}} f(A) \geq \delta\left(\bigcup_{A \in \mathcal{A}} A\right).$$

If $A_0 \in \mathcal{A}$ then

$$\begin{aligned}
\sum_{A \in \mathcal{A}} g(A) &= \sup_{\mathcal{B} \subseteq \mathcal{P}_{\text{fin}}(X)} \left\{ \delta(A_0 \cup \bigcup_{B \in \mathcal{B}} B) - \sum_{B \in \mathcal{B}} f(B) : |\mathcal{B}| < \infty \right\} + \sum_{B \in \mathcal{A} \setminus \{A_0\}} f(B) \\
&\geq \delta(A_0 \cup \bigcup_{B \in \mathcal{A} \setminus \{A_0\}} B) \\
&= \delta(\bigcup_{A \in \mathcal{A}} A),
\end{aligned}$$

by letting $\mathcal{B} = \mathcal{A} \setminus \{A_0\}$. So $g \in P_X$, $g \neq f$ and $g \preceq f$, contradicting $f \in T_X$. Hence there is no A_0 satisfying (2.4). If $f \in T_X$ then (2.3) holds for all finite $A \subseteq X$.

For the converse, suppose that (2.3) holds for all finite $A \subseteq X$. Then $f \in P_X$. Suppose that $g \in P_X$, that $g \preceq f$ and $A \in \mathcal{P}_{\text{fin}}(X)$. Then for all finite $\mathcal{B} \subseteq \mathcal{P}_{\text{fin}}(X)$ we have

$$\delta(A \cup \bigcup_{B \in \mathcal{B}} B) - \sum_{B \in \mathcal{B}} f(B) \leq \delta(A \cup \bigcup_{B \in \mathcal{B}} B) - \sum_{B \in \mathcal{B}} g(B) \leq g(A)$$

so that $f(A) \leq g(A)$. Hence f is minimal in P_X . \square

We note that the characterisation of tight spans given by Theorem 2.3 is analogous to the definition of tight spans for valuated matroids used by [31]. One important difference is that, for diversities, the tight span is made up of functions on $\mathcal{P}_{\text{fin}}(X)$ rather than functions on X .

The following basic properties of members of T_X will be used subsequently.

Proposition 2.4. *Suppose that $f \in T_X$.*

1. $f(A) \geq \delta(A)$ for all finite $A \subseteq X$.
2. If $A \subseteq B \subseteq X$ and B is finite then $f(A) \leq f(B)$; that is, f is monotone.
3. $f(A \cup C) \leq \delta(A \cup B) + f(B \cup C)$ for all $A, B, C \in \mathcal{P}_{\text{fin}}(X)$ with $B \neq \emptyset$.
4. $f(A \cup B) \leq f(A) + f(B)$ for all $A, B \in \mathcal{P}_{\text{fin}}(X)$; that is, f is sub-additive.
5. $f(A) = \sup_B \{\delta(A \cup B) - f(B) : B \in \mathcal{P}_{\text{fin}}(X)\}$ for all finite A .

Proof.

1. Use $\mathcal{A} = \{A\}$ in the definition of P_X .

2. Follows from (2.3) and the monotonicity of δ .

3. Let $A, B, C \in \mathcal{P}_{\text{fin}}(X)$ with $B \neq \emptyset$. We have

$$f(A \cup C) = \sup_{\mathcal{D} \subseteq \mathcal{P}_{\text{fin}}(X)} \left\{ \delta(A \cup C \cup \bigcup_{D \in \mathcal{D}} D) - \sum_{D \in \mathcal{D}} f(D) : |\mathcal{D}| < \infty \right\} \quad (2.5)$$

$$f(B \cup C) = \sup_{\mathcal{D} \subseteq \mathcal{P}_{\text{fin}}(X)} \left\{ \delta(B \cup C \cup \bigcup_{D \in \mathcal{D}} D) - \sum_{D \in \mathcal{D}} f(D) : |\mathcal{D}| < \infty \right\}. \quad (2.6)$$

Subtracting (2.6) from (2.5) gives

$$\begin{aligned} f(A \cup C) - f(B \cup C) &\leq \sup_{\mathcal{D} \subseteq \mathcal{P}_{\text{fin}}(X)} \left\{ \delta(A \cup C \cup \bigcup_{D \in \mathcal{D}} D) - \delta(B \cup C \cup \bigcup_{D \in \mathcal{D}} D) : |\mathcal{D}| < \infty \right\} \\ &\leq \delta(A \cup B), \end{aligned}$$

by taking $\mathcal{D} = \emptyset$ and using the triangle inequality. We note that this property is analogous to the continuity of functions in the metric tight span, see [4, Theorem 3(iv)].

4. Given any $A, B \in \mathcal{P}_{\text{fin}}(X)$ and any finite collection $\mathcal{C} \subseteq \mathcal{P}_{\text{fin}}(X)$ we have

$$f(A) + f(B) + \sum_{C \in \mathcal{C}} f(C) \geq \delta(A \cup B \cup \bigcup_{C \in \mathcal{C}} C)$$

so that

$$\begin{aligned} f(A) + f(B) &\geq \sup_{\mathcal{C} \subseteq \mathcal{P}_{\text{fin}}(X)} \left\{ \delta(A \cup B \cup \bigcup_{C \in \mathcal{C}} C) - \sum_{C \in \mathcal{C}} f(C) : |\mathcal{C}| < \infty \right\} \\ &= f(A \cup B) \end{aligned}$$

by Theorem 2.3.

5. For any finite $\mathcal{B} \subseteq \mathcal{P}_{\text{fin}}$, $\sum_{B \in \mathcal{B}} f(A) \geq f(\cup_{B \in \mathcal{B}} B)$. So

$$\sup_{\mathcal{B} \in \mathcal{P}_{\text{fin}}(X)} \left\{ \delta(A \cup \bigcup_{B \in \mathcal{B}} B) - \sum_{B \in \mathcal{B}} f(B) : |\mathcal{B}| < \infty \right\} = \sup_{C \in \mathcal{P}_{\text{fin}}(X)} \{ \delta(A \cup C) - f(C) \}.$$

□

The distance between any two functions f, g in the metric tight span is given by the l_∞ norm,

$$d_T(f, g) = \sup_{x \in X} |f(x) - g(x)| \quad (2.7)$$

which Dress [4, Theorem 3(iii)] shows is equivalent on this set to

$$d_T(f, g) = \sup_{x, y \in X} \{d(x, y) - f(x) - g(y)\}. \quad (2.8)$$

Dress also showed that a metric can be embedded into its tight span using the Kuratowski map κ , which takes an element $x \in X$ to the function h_x for which $h_x(y) = d(x, y)$ for all y . This is exactly the map e defined in [3, section 2.4].

Here we establish the analogous results for the diversity tight span. We define the appropriate function κ from a diversity to its tight span. We then define a function δ_T on T_X so that (T_X, δ_T) is a diversity and prove that κ is an embedding.

- Definition 2.5.** 1. Let (Y_1, δ_1) and (Y_2, δ_2) be two diversities. A map $\pi: Y_1 \rightarrow Y_2$ is an embedding if it is one-to-one (injective) and for all finite $A \subseteq Y_1$ we have $\delta_1(A) = \delta_2(\pi(A))$. In this case, we say that π embeds (Y_1, δ_1) in (Y_2, δ_2) .
2. An isomorphism is an onto (surjective) embedding between two diversities.
3. Let (X, δ) be a diversity. For each $x \in X$ define the function $h_x : \mathcal{P}_{\text{fin}}(X) \rightarrow \mathbb{R}$ by

$$h_x(A) = \delta(A \cup \{x\})$$

for all finite $A \subseteq X$. Let κ be the map taking each $x \in X$ to the corresponding function h_x .

4. Let (X, δ) be a diversity. Let $\delta_T : \mathcal{P}_{\text{fin}}(T_X) \rightarrow \mathbb{R}$ be the function defined by $\delta_T(\emptyset) = 0$ and

$$\delta_T(F) = \sup_{\mathcal{A} \subseteq \mathcal{P}_{\text{fin}}(X)} \left\{ \delta \left(\bigcup_{A \in \mathcal{A}} A \right) - \sum_{A \in \mathcal{A}} \inf_{f \in F} f(A) : |\mathcal{A}| < \infty \right\} \quad (2.9)$$

for all finite non-empty $F \subseteq T_X$.

Further manipulations give a form for δ_T analogous to (2.8):

$$\delta_T(F) = \sup_{\{A_f\}_{f \in F}} \left\{ \delta \left(\bigcup_{f \in F} A_f \right) - \sum_{f \in F} f(A_f) : A_f \in \mathcal{P}_{\text{fin}}(X) \text{ for all } f \in F \right\},$$

for all finite $F \subseteq \mathcal{P}_{\text{fin}}(T_X)$. We can also re-express (2.9) in a form closer to (2.7).

Note the similarity between Lemma 2.6 and [4, Theorem 3(iii)].

Lemma 2.6. *If $f \in F$ then*

$$\delta_T(F) = \sup_{\mathcal{A} \subseteq \mathcal{P}_{\text{fin}}(X)} \left\{ f \left(\bigcup_{A \in \mathcal{A}} A \right) - \sum_{A \in \mathcal{A}} \inf_{g \in F \setminus \{f\}} g(A) : |\mathcal{A}| < \infty \right\}.$$

Proof.

$$\begin{aligned} \delta_T(F) &= \sup_{\mathcal{A}, \mathcal{B} \subseteq \mathcal{P}_{\text{fin}}(X)} \left\{ \delta \left(\bigcup_{A \in \mathcal{A}} A \cup \bigcup_{B \in \mathcal{B}} B \right) - \sum_{A \in \mathcal{A}} \inf_{g \in F \setminus \{f\}} g(A) - \sum_{B \in \mathcal{B}} f(B) : |\mathcal{A}|, |\mathcal{B}| < \infty \right\} \\ &= \sup_{\mathcal{A} \subseteq \mathcal{P}_{\text{fin}}(X)} \left\{ f \left(\bigcup_{A \in \mathcal{A}} A \right) - \sum_{A \in \mathcal{A}} \inf_{g \in F \setminus \{f\}} g(A) : |\mathcal{A}| < \infty \right\}, \end{aligned}$$

by Theorem 2.3. □

Theorem 2.7. *(T_X, δ_T) is a diversity.*

Proof.

First note that for all $F \subseteq T_X$, when $\mathcal{A} = \{\emptyset\}$,

$$\delta \left(\bigcup_{A \in \mathcal{A}} A \right) - \sum_{A \in \mathcal{A}} \inf_{f \in F} f(A) = 0$$

so that δ_T is non-negative.

If $\emptyset \neq F \subseteq G$ then for all $\mathcal{A} \subseteq \mathcal{P}_{\text{fin}}(X)$ with $|\mathcal{A}| < \infty$ we have

$$\sum_{A \in \mathcal{A}} \inf_{f \in F} f(A) \geq \sum_{A \in \mathcal{A}} \inf_{f \in G} f(A).$$

Hence $\delta_T(F) \leq \delta_T(G)$, showing that δ_T is monotone.

If $F = \{f\}$ then

$$\delta_T(F) = \delta_T(\{f\}) \leq \sup \{ \delta(A) - f(A) : A \in \mathcal{P}_{\text{fin}}(X) \} = 0$$

by the subadditivity of f and by part 1 of Proposition 2.4. On the other hand, if $|F| > 1$ then there is $f_1, f_2 \in F$ such that $f_1 \neq f_2$. By monotonicity and

Lemma 2.6 we have

$$\begin{aligned}
\delta_T(F) &\geq \delta_T(\{f_1, f_2\}) \\
&= \sup_{\mathcal{A} \subseteq \mathcal{P}_{\text{fin}}(X)} \left\{ f_1 \left(\bigcup_{A \in \mathcal{A}} A \right) - f_2 \left(\bigcup_{A \in \mathcal{A}} A \right) : |\mathcal{A}| < \infty \right\} \\
&= \sup_{A \in \mathcal{P}_{\text{fin}}(X)} \{f_1(A) - f_2(A)\} \\
&> 0.
\end{aligned}$$

We have now proved that δ_T satisfies (D1).

For the triangle inequality, suppose F and G are disjoint finite subsets of T_X and that $h \in T_X \setminus (F \cup G)$. Then by Lemma 2.6

$$\delta_T(F \cup \{h\}) = \sup_{\mathcal{A} \subseteq \mathcal{P}_{\text{fin}}(X)} \left\{ h \left(\bigcup_{A \in \mathcal{A}} A \right) - \sum_{A \in \mathcal{A}} \inf_{f \in F} f(A) : |\mathcal{A}| < \infty \right\} \quad (2.10)$$

and

$$\delta_T(G \cup \{h\}) = \sup_{\mathcal{B} \subseteq \mathcal{P}_{\text{fin}}(X)} \left\{ h \left(\bigcup_{B \in \mathcal{B}} B \right) - \sum_{B \in \mathcal{B}} \inf_{g \in G} g(B) : |\mathcal{B}| < \infty \right\} \quad (2.11)$$

By part 4 of Proposition 2.4 the function h is sub-additive, so

$$h \left(\bigcup_{A \in \mathcal{A}} A \right) + h \left(\bigcup_{B \in \mathcal{B}} B \right) \geq h \left(\bigcup_{C \in \mathcal{A} \cup \mathcal{B}} C \right). \quad (2.12)$$

Combining (2.10)–(2.12) and again applying Lemma 2.6 we have

$$\begin{aligned}
\delta_T(F \cup \{h\}) + \delta_T(G \cup \{h\}) &\geq \sup_{\mathcal{C} \subseteq \mathcal{P}_{\text{fin}}(X)} \left\{ h \left(\bigcup_{C \in \mathcal{C}} C \right) - \sum_{C \in \mathcal{C}} \inf_{f \in F \cup G} f(C) : |\mathcal{C}| < \infty \right\} \\
&= \delta_T(F \cup G \cup \{h\}).
\end{aligned}$$

The triangle inequality (D2) now follows by monotonicity. \square

Theorem 2.7 establishes that (T_X, δ_T) is a diversity. We now show that κ is an embedding from (X, δ) into (T_X, δ_T) . We then prove the diversity analogue of [3, Eq. (2.4)] (see [4, Theorem 3(ii)]) and characterise δ_T in terms of a minimality condition.

Theorem 2.8. 1. The map κ is an embedding from (X, δ) into (T_X, δ_T) .
2. For all finite $Y \subseteq X$ and $f \in T_X$,

$$\delta_T(\kappa(Y) \cup \{f\}) = f(Y).$$

3. If $(T_X, \widehat{\delta})$ is a diversity such that $\widehat{\delta}(\kappa(Y) \cup \{f\}) = f(Y)$ for all finite $Y \subseteq X$ and $f \in T_X$ then

$$\widehat{\delta}(F) \geq \delta_T(F)$$

for all finite $F \subseteq T_X$.

Proof.

1. Fix $x \in X$. Consider finite $\mathcal{A} \subseteq \mathcal{P}_{\text{fin}}(X)$. The triangle inequality for diversities, (D2), gives

$$\sum_{A \in \mathcal{A}} h_x(A) = \sum_{A \in \mathcal{A}} \delta(A \cup \{x\}) \geq \delta\left(\bigcup_{A \in \mathcal{A}} A\right),$$

so that $h_x \in P_X$. There is $g \in T_X$ such that $g \preceq h_x$. Since $h_x(\{x\}) = \delta(\{x\}) = 0$ we have for all finite $A \subseteq X$ that

$$h_x(A) = \delta(A \cup \{x\}) \leq g(A) + g(\{x\}) \leq g(A) + h_x(\{x\}) = g(A) \leq h_x(A).$$

Hence $h_x = g \in T_X$.

To see that κ is one-to-one observe that for $x \neq y$, $h_x(\{x\}) = 0$ but $h_y(\{x\}) = \delta(\{x, y\}) > 0$. So $h_x \neq h_y$ for distinct $x, y \in X$.

We now show that $\delta_T(\kappa(Y)) = \delta(Y)$ for all finite $Y \subseteq X$. Let $Y \subseteq X$, $Y = \{y_1, \dots, y_k\}$. Taking $\mathcal{A} = \{\{y_1\}, \dots, \{y_k\}\}$ in (2.9) gives $\delta_T(\kappa(Y)) \geq \delta(Y)$.

By repeatedly using the triangle inequality we have for any finite $\mathcal{A} =$

$\{A_1, A_2, \dots, A_j\} \subseteq \mathcal{P}_{\text{fin}}(X)$ and $z_1, \dots, z_j \in Y$ that

$$\begin{aligned}
\delta(Y) &\geq \delta(Y \cup A_1) - \delta(\{z_1\} \cup A_1) \\
&\geq \delta(Y \cup A_1 \cup A_2) - \delta(\{z_1\} \cup A_1) - \delta(\{z_2\} \cup A_2) \\
&\geq \delta\left(Y \cup \bigcup_{i=1}^j A_i\right) - \sum_{i=1}^j \delta(\{z_i\} \cup A_i) \\
&\geq \delta\left(\bigcup_{i=1}^j A_i\right) - \sum_{i=1}^j h_{z_i}(A_i) \\
&\geq \delta\left(\bigcup_{i=1}^j A_i\right) - \sum_{i=1}^j \inf_{h \in \kappa(Y)} h(A_i).
\end{aligned}$$

Taking the supremum over all such \mathcal{A} and applying (2.9) gives $\delta_T(\kappa(Y)) \leq \delta(Y)$.

So $\delta_T(\kappa(Y)) = \delta(Y)$ and κ is an embedding.

2. Let $Y \subseteq X$, Y finite, and $f \in T_X$. If $f = h_y$ for $y \in Y$ then, using part 1,

$$\delta_T(\kappa(Y) \cup \{f\}) = \delta_T(\kappa(Y)) = \delta(Y) = \delta(Y \cup \{y\}) = f(Y),$$

as required. Otherwise, suppose $f \notin \kappa(Y)$. Let $Y = \{y_1, \dots, y_k\}$.

$$\delta_T(\kappa(Y) \cup \{f\}) = \sup_{A_i, i=1, \dots, k, A_f} \left\{ \delta\left(\bigcup_i A_i \cup A_f\right) - \sum_i \delta(\{y_i\} \cup A_i) - f(A_f) \right\}.$$

Letting $A_i = \{y_i\}$ for all i shows

$$\delta_T(\kappa(Y) \cup \{f\}) \geq \sup_{A_f} \{\delta(Y \cup A_f) - f(A_f)\} = f(Y),$$

by Proposition 2.4 part 5. On the other hand, following the same reasoning as in part 1 of this proof shows

$$\delta_T(\kappa(Y) \cup \{f\}) \leq \sup_{A_f} \delta(Y \cup A_f) - f(A_f) = f(Y).$$

Therefore $\delta_T(\kappa(Y) \cup \{f\}) = f(Y)$.

3. Suppose that $F = \kappa(Y) \cup G$, where $Y \in \mathcal{P}_{\text{fin}}(X)$ and $G \subseteq T_X \setminus \kappa(X)$.

For all collections $\mathcal{A} \subseteq \mathcal{P}_{\text{fin}}(X)$ with $|\mathcal{A}| < \infty$ and all collections $\{f_A\}_{A \in \mathcal{A}}$ of

elements in F , we have from 1. and 2. that

$$\begin{aligned}
\delta\left(Y \cup \bigcup_{A \in \mathcal{A}} A\right) - \sum_{A \in \mathcal{A}} f_A(A) &= \widehat{\delta}\left(\kappa(Y) \cup \bigcup_{A \in \mathcal{A}} \kappa(A)\right) - \sum_{A \in \mathcal{A}} \widehat{\delta}(\kappa(A) \cup \{f_A\}) \\
&\leq \widehat{\delta}(\kappa(Y) \cup \{f_A : A \in \mathcal{A}\}) \\
&\leq \widehat{\delta}(\kappa(Y) \cup F).
\end{aligned}$$

□

3. Hyperconvex diversities and the injective envelope

Aronszajn and Panitchpakdi [1] introduced hyperconvex metric spaces and showed that they are exactly the injective metric spaces.

Definition 3.1. 1. A metric space (X, d) is said to be hyperconvex if for all $r : X \rightarrow \mathbb{R}$ with $r(x) + r(y) \geq d(x, y)$ for all $x, y \in X$ there is a point $z \in X$ such that $d(z, x) \leq r(x)$ for all $x \in X$.

2. A metric space (X, d) is injective if it satisfies the following property: given any pair of metric spaces (Y_1, d_1) , (Y_2, d_2) , an embedding $\pi : Y_1 \rightarrow Y_2$ and a non-expansive map $\phi : Y_1 \rightarrow X$ there is a non-expansive map $\psi : Y_2 \rightarrow X$ such that $\phi = \psi \circ \pi$.

See [2] for a proof of the equivalence of these two concepts, as well as a highly readable and comprehensive review of the rich metric structure of hyperconvex spaces. Here we establish diversity analogues for these concepts and show that the equivalence holds in this new setting. We begin by defining diversity analogues of injective and hyperconvex metric spaces.

Definition 3.2. 1. Given diversities (Y_1, δ_1) and (Y_2, δ_2) , a map $\phi : Y_1 \rightarrow Y_2$ is non-expansive if for all $A \subseteq Y_1$ we have $\delta_1(A) \geq \delta_2(\phi(A))$ and it is an embedding if it is one-to-one and for all $A \subseteq Y_1$ we have $\delta_1(A) = \delta_2(\phi(A))$.

2. A diversity (X, δ) is injective if it satisfies the following property: given any pair of diversities (Y_1, δ_1) , (Y_2, δ_2) , an embedding $\pi : Y_1 \rightarrow Y_2$ and a non-expansive map $\phi : Y_1 \rightarrow X$ there is a non-expansive map $\psi : Y_2 \rightarrow X$ such that $\phi = \psi \circ \pi$.
3. A diversity (X, δ) is said to be hyperconvex if for all $r : \mathcal{P}_{\text{fin}}(X) \rightarrow \mathfrak{R}$ such that

$$\delta \left(\bigcup_{A \in \mathcal{A}} A \right) \leq \sum_{A \in \mathcal{A}} r(A) \quad (3.1)$$

for all finite $\mathcal{A} \subseteq \mathcal{P}_{\text{fin}}(X)$ there is $z \in X$ such that $\delta(\{z\} \cup Y) \leq r(Y)$ for all finite $Y \subseteq X$.

The following theorem establishes the diversity equivalent of Aronszajn and Panitchpakdi's result.

Theorem 3.3. *A diversity (X, δ) is injective if and only if it is hyperconvex.*

Proof.

First suppose that (X, δ) is injective. Consider $r : \mathcal{P}_{\text{fin}}(X) \rightarrow \mathfrak{R}$ satisfying (3.1) for all finite $\mathcal{A} \subseteq \mathcal{P}_{\text{fin}}(X)$. Without loss of generality we can assume $r(\emptyset) = 0$ and hence $r \in P_X$. Choose $f \in T_X$ with $f \preceq r$.

Let x^* be a point not in X , let $X^* = X \cup \{x^*\}$ and let $\delta^* : \mathcal{P}_{\text{fin}}(X \cup \{x^*\}) \rightarrow \mathfrak{R}$ be the function where for all finite $A \subseteq X$,

$$\begin{aligned} \delta^*(A) &= \delta(A) \\ \delta^*(A \cup \{x^*\}) &= f(A). \end{aligned}$$

From Proposition 2.4 part 2 we have that δ^* is monotonic, and from part 4 and 5 we have that

$$\delta^*(A \cup C \cup \{x^*\}) \leq \delta^*(A \cup \{x^*\}) + \delta^*(C \cup \{x^*\}) \quad (3.2)$$

$$\delta^*(A \cup B \cup C \cup \{x^*\}) \leq \delta^*(A \cup B \cup \{x^*\}) + \delta^*(B \cup C). \quad (3.3)$$

for all finite $A, B, C \subseteq X$ such that $B \neq \emptyset$. These, together with monotonicity and the fact that δ^* coincides with δ on $\mathcal{P}_{\text{fin}}(X)$, imply the triangle inequality (D2) for (X^*, δ^*) .

We now apply the fact that (X, δ) is injective. Let (Y_1, δ_1) be (X, δ) ; let (Y_2, δ_2) be (X^*, δ^*) , let π be the identity embedding from (X, δ) into (X^*, δ^*) and let ϕ be the identity map from (X, δ) to itself. Then there is a non-expansive map $\phi : X^* \rightarrow X$ such that $\phi(x) = x$ for all $x \in X$.

Let $\omega = \phi(x^*)$. For all finite $A \subseteq X$ we have

$$\begin{aligned} \delta(A \cup \{\omega\}) &\leq \delta^*(A \cup \{x^*\}) \\ &= f(A) \\ &\leq r(A). \end{aligned}$$

This proves that (X, δ) is hyperconvex.

For the converse, suppose now that (X, δ) is hyperconvex. Let (Y_1, δ_1) and (Y_2, δ_2) be two diversities, let $\pi : Y_1 \rightarrow Y_2$ be an embedding and let ϕ be a non-expansive map from Y_1 to X . We will show that there is non-expansive $\psi : Y_2 \rightarrow X$ such that $\phi = \psi \circ \pi$.

Let \mathcal{Y} denote the collection of pairs (Y, ψ_Y) such that $\pi(Y_1) \subseteq Y \subseteq Y_2$ and ψ_Y is a non-expansive map from Y to X such that $\phi = \psi_Y \circ \pi$. We want to show that $Y_2 \in \mathcal{Y}$. Suppose this is not the case. We write $(Y, \psi_Y) \trianglelefteq (Z, \psi_Z)$ if $Y \subseteq Z$ and ψ_Z restricted to Y equals ψ_Y . The partially ordered set $(\mathcal{Y}, \trianglelefteq)$ satisfies the conditions of Zorn's lemma, so it contains maximal elements.

Let (Y, ψ_Y) be one such maximal element. Choose $y \in Y_2 \setminus Y$. For each finite $A \subseteq Y$ let $r(A) = \delta_2(A \cup \{y\})$. For any finite collection $\mathcal{A} \subseteq \mathcal{P}_{\text{fin}}(Y)$ we have

$$\begin{aligned} \delta\left(\bigcup_{A \in \mathcal{A}} \psi_Y(A)\right) &= \delta\left(\psi_Y\left(\bigcup_{A \in \mathcal{A}} A\right)\right) \\ &\leq \delta_2\left(\bigcup_{A \in \mathcal{A}} A\right) \\ &\leq \sum_{A \in \mathcal{A}} \delta_2(A \cup \{y\}) \\ &= \sum_{A \in \mathcal{A}} r(A). \end{aligned}$$

Since (X, δ) is hyperconvex, there is $x \in X$ such that

$$\delta(\psi_Y(A) \cup \{x\}) \leq r(A) = \delta_2(A \cup \{y\})$$

for all finite $A \subseteq Y$. Hence we can extend ψ_Y to $Y \cup \{y\}$ by setting $\psi_Y(y) = x$, giving a non-expansive map from $Y \cup \{y\}$ to X , and contradicting the maximality of Y .

It follows that $Y_2 \in \mathcal{Y}$, proving that (X, δ) is injective. \square

Definition 3.4. Let (X, δ) be a diversity. For $F \subseteq T_X$ and finite $Y \subseteq X$ let

$$\Phi_F(Y) = \inf_{\mathcal{A} \subseteq \mathcal{P}_{\text{fin}}(X)} \left\{ \sum_{A \in \mathcal{A}} \inf_{f \in F} f(A) : |\mathcal{A}| < \infty, \bigcup_{A \in \mathcal{A}} A = Y \right\}.$$

Clearly,

$$\delta_T(F) = \sup_{Y \subseteq X} \{\delta(Y) - \Phi_F(Y) : |Y| < \infty\}. \quad (3.4)$$

We show that Φ_F also satisfies a sub-additivity type identity.

Lemma 3.5. For $F, G \subseteq T_X$ and $Y, Z \subseteq \mathcal{P}_{\text{fin}}(X)$ we have

$$\Phi_{F \cup G}(Y \cup Z) \leq \Phi_F(Y) + \Phi_G(Z).$$

Proof.

Given $\epsilon > 0$ there is finite $\mathcal{A} \subseteq \mathcal{P}_{\text{fin}}(X)$ and a collection $\{f_A\}_{A \in \mathcal{A}}$ of elements in T_X such that

$$\Phi_F(Y) \leq \sum_{A \in \mathcal{A}} f_A(A) < \Phi_F(Y) + \epsilon/2.$$

Similarly, there is finite $\mathcal{B} \subseteq \mathcal{P}_{\text{fin}}(X)$ and a collection $\{g_B\}_{B \in \mathcal{B}}$ of elements in T_X such that

$$\Phi_G(Z) \leq \sum_{B \in \mathcal{B}} g_B(B) < \Phi_G(Z) + \epsilon/2.$$

Define $\mathcal{C} = \mathcal{A} \cup \mathcal{B}$ and the collection $\{h_C\}_{C \in \mathcal{C}}$ by

$$h_C = \begin{cases} f_C & \text{if } C \in \mathcal{A}; \\ g_C & \text{otherwise.} \end{cases}$$

Then

$$\begin{aligned}
\Phi_F(Y) + \Phi_G(Z) + \epsilon &> \sum_{A \in \mathcal{A}} f_A(A) + \sum_{B \in \mathcal{B}} g_B(B) \\
&\geq \sum_{C \in \mathcal{C}} h_C(C) \\
&\geq \Phi_{F \cup G}(Y \cup Z).
\end{aligned}$$

Taking $\epsilon \rightarrow 0$ proves the Lemma. \square

Isbell proved that the metric tight span is injective, and hence hyperconvex [3, Section 2.9]. Here we prove the same result for diversities.

Theorem 3.6. *For any diversity (X, δ) , the tight span (T_X, δ_T) is hyperconvex.*

Proof.

Let $r: \mathcal{P}_{\text{fin}}(T_X) \rightarrow \mathfrak{R}$ be given such that for all finite $\mathcal{F} \subseteq \mathcal{P}_{\text{fin}}(T_X)$

$$\sum_{F \in \mathcal{F}} r(F) \geq \delta_T \left(\bigcup_{F \in \mathcal{F}} F \right).$$

Without loss of generality we can assume $r(\emptyset) = 0$. We need to find $g \in T_X$ so that $\delta_T(G \cup \{g\}) \leq r(G)$ for all $G \subseteq T_X$.

Define ω on $\mathcal{P}_{\text{fin}}(X)$ by

$$\omega(A) = \inf_{F \subseteq T_X} \{r(F) + \Phi_F(A) : |F| < \infty\}.$$

We have $\omega(\emptyset) = 0$. Suppose that $\mathcal{A} \subseteq \mathcal{P}_{\text{fin}}(X)$, $|\mathcal{A}| < \infty$ and let $\{F_A : A \in \mathcal{A}\}$ be a collection of finite subsets of T_X indexed by elements of \mathcal{A} . From Lemma 3.5 and (3.4) we have

$$\begin{aligned}
\delta \left(\bigcup_{A \in \mathcal{A}} A \right) &\leq \delta_T \left(\bigcup_{A \in \mathcal{A}} F_A \right) + \Phi_{(\bigcup_{A \in \mathcal{A}} F_A)} \left(\bigcup_{A \in \mathcal{A}} A \right) \\
&\leq \sum_{A \in \mathcal{A}} (r(F_A) + \Phi_{F_A}(A)),
\end{aligned}$$

so that

$$\delta \left(\bigcup_{A \in \mathcal{A}} A \right) \leq \sum_{A \in \mathcal{A}} \omega(A),$$

and $\omega \in P_X$.

There is $g \in T_X$ such that $g \preceq \omega$. Consider finite $F \subseteq T_X$. Applying Lemma 2.6,

$$\begin{aligned} \delta_T(F \cup \{g\}) &= \sup_{A \in \mathcal{P}_{\text{fin}}(X)} \{g(A) - \Phi_F(A)\} \\ &\leq \sup_{A \in \mathcal{P}_{\text{fin}}(X)} \{(r(F) + \Phi_F(A)) - \Phi_F(A)\} \\ &= r(F), \end{aligned}$$

as required. \square

The metric tight span construction gives an isometric embedding κ from a metric space (X, d) into an injective (hyperconvex) metric space. Isbell showed that this embedding is minimal in that no proper subspace of the tight span both contains $\kappa(X)$ and is injective. Such an embedding is called an *injective envelope*, and all injective envelopes of a metric space are equivalent [3, Thm 2.1].

Here we prove the analogous result for diversities that the embedding κ of a diversity into its tight span is also an injective envelope.

The class of all diversities with all non-expansive maps as morphisms forms a category, which we will denote **Dvy**. The definitions of embeddings and injective objects then correspond to concepts in category theory, as reviewed in [33]. Lemma 3.7 together with the injectivity of (T_X, δ_T) establishes that (T_X, δ_T) is the *injective hull* of (X, δ) in the category **Dvy** [33, pg. 156]. Proposition 9.20(5) of [33] demonstrates the equivalence between the category theory injective hull and the injective envelope introduced in [3].

Lemma 3.7. *Let ϕ be a non-expanding map from (T_X, δ_T) to diversity (Y, δ_Y) . If $\pi = \phi \circ \kappa$ is an embedding from (X, δ) to (Y, δ_Y) then ϕ is an embedding from (T_X, δ_T) to (Y, δ_Y) .*

Proof.

Since ϕ is non-expanding $\delta_T(F) \geq \delta_Y(\phi(F))$ for all finite $F \subseteq T_X$. Using

Theorem 2.8 part 3 we will show that $\delta_T(F) \leq \delta_Y(\phi(F))$, so that ϕ is an embedding.

Consider $f \in T_X$. Define g on $\mathcal{P}_{\text{fin}}(X)$ by $g(A) = \delta_Y(\pi(A) \cup \phi(\{f\}))$ for all finite A . Then for any finite $A \subseteq X$ we have

$$g(A) = \delta_Y(\pi(A) \cup \phi(\{f\})) = \delta_Y(\phi(\kappa(A) \cup \{f\})) \leq \delta_T(\kappa(A) \cup \{f\}) = f(A)$$

for all A . For all finite collections $\mathcal{A} \subseteq \mathcal{P}_{\text{fin}}(X)$ we have

$$\begin{aligned} \sum_{A \in \mathcal{A}} g(A) &= \sum_{A \in \mathcal{A}} \delta_Y(\pi(A) \cup \phi(\{f\})) \\ &\geq \delta_Y\left(\bigcup_{A \in \mathcal{A}} \pi(A)\right) \\ &= \delta\left(\bigcup_{A \in \mathcal{A}} A\right), \end{aligned}$$

so that $g \in P_X$ and $g \preceq f$. Hence $g(A) = f(A)$ for all finite $A \subseteq X$. It follows that

$$\delta_Y(\pi(A) \cup \phi(\{f\})) = \delta_T(\kappa(A) \cup \{f\})$$

for all $f \in T_X$ and finite $A \subseteq X$.

Define $\widehat{\delta}$ on T_X by $\widehat{\delta}(F) = \delta_Y(\phi(F))$. Then $\widehat{\delta}$ is a diversity and $\widehat{\delta}(\kappa(Y) \cup \{f\}) = f(Y)$ for all finite $Y \subseteq X$. By Theorem 2.8, $\widehat{\delta}(F) \geq \delta_T(F)$ for all finite F .

As ϕ is non-expansive $\delta_T(F) = \delta_Y(\phi(F))$ for all finite F and ϕ is an embedding. \square

The following Theorem is a translation of [33, Proposition 9.20(4)] to diversities.

Theorem 3.8. *If there is an embedding π from (X, δ) into (Y, δ_Y) and (Y, δ_Y) is injective (hyperconvex) then there is an embedding ϕ from (T_X, δ_T) into (Y, δ_Y) such that $\pi = \phi \circ \kappa$.*

Proof.

Since π is a non-expansive map, (Y, δ_Y) is injective, and κ is an embedding

of (X, δ) into (T_X, δ_T) , there is a non-expansive map $\phi : T_X \rightarrow Y$ such that $\pi = \phi \circ \kappa$. By Lemma 3.7, ϕ is an embedding. \square

Corollary 3.9. *Let (X, δ) be a diversity. The following are equivalent:*

1. (X, δ) is hyperconvex;
2. (X, δ) is injective;
3. There is an isomorphism between (X, δ) and its tight span, (T_X, δ_T) .

Proof.

1. and 2. are equivalent by Theorem 3.3. To see that 2. implies 3., let $(Y, \delta_Y) = (X, \delta)$ and $\pi = \text{id}$ in Theorem 3.8. Then there is an embedding ϕ from (T_X, δ_T) to (X, δ) such that $\phi \circ \kappa = \text{id}$. So κ is surjective and 3. follows. Finally, since hyperconvexity is invariant under isomorphism, 1. follows from 3. \square

4. Tight span of the diameter diversity

In this section we prove that tight span theory for metrics is embedded within the tight span theory for diversities. The link between the two is provided by the *diameter diversity* as introduced above.

Definition 4.1. *Given a metric space (X, d) we define the function $\delta = \text{diam}_d$ by*

$$\delta(A) = \text{diam}_d(A) = \max\{d(a, a') : a, a' \in A\}$$

for finite $A \subseteq X$, with $\text{diam}_d(\emptyset) = 0$. We call (X, diam_d) the diameter diversity for (X, d) . Note that if we restrict diam_d to pairs of elements we recover d as the induced metric.

We will establish close links between tight spans of metrics and tight spans of their diameter diversities.

$$\begin{array}{ccc}
(X, d) & \xrightarrow{\text{tight span}} & (T_X^d, d_T) \\
\delta = \text{diam}_d \downarrow & & \downarrow \delta_T = \text{diam}_{d_T} \\
(X, \delta) & \xrightarrow{\text{tight span}} & (T_X^\delta, \delta_T)
\end{array}$$

- Lemma 4.2.** 1. Let (Y, δ) be a diversity with induced metric (Y, d_δ) . Let (X, d) be a metric space and let (X, diam_d) be the associated diameter diversity. Then ϕ is a non-expansive map from (Y, δ) to (X, diam_d) if and only if it is a non-expansive map from (Y, d_δ) to (X, d) .
2. A metric space (X, d) is injective (hyperconvex) if and only if the diameter diversity (X, diam_d) is injective (hyperconvex).
3. The tight span (T_X^δ, δ_T) of a diameter diversity is itself a diameter diversity.

Proof.

1. Suppose that ϕ is a non-expansive map from (Y, δ) to (X, diam_d) . For all $y_1, y_2 \in Y$ we have

$$d_\delta(y_1, y_2) = \delta(\{y_1, y_2\}) \geq \text{diam}_d(\{\phi(y_1), \phi(y_2)\}) = d(\phi(y_1), \phi(y_2)),$$

so ϕ is non-expansive from (Y, d_δ) to (X, d) . Conversely, suppose ϕ is a non-expansive map from (Y, d_δ) to (X, d) . Then for any finite $A \subseteq Y$ we have

$$\begin{aligned}
\delta(A) &\geq \sup\{d_\delta(a_1, a_2) : a_1, a_2 \in A\} \\
&\geq \sup\{d(\phi(a_1), \phi(a_2)) : a_1, a_2 \in A\} \\
&= \text{diam}_d(\phi(A)).
\end{aligned}$$

2. Suppose that (X, d) is injective. Let $(Y_1, \delta_1), (Y_2, \delta_2)$ be two diversities with induced metrics d_1, d_2 . Let π be an embedding from (Y_1, δ_1) into (Y_2, δ_2) and let ϕ be a non-expansive map from (Y_1, δ_1) to (X, diam_d) . Then π embeds (Y_1, d_1) into (Y_2, d_2) , and by part 1., ϕ is a non-expansive map from (Y_1, d_1) to (X, d) . As (X, d) is an injective metric space there is a non-expansive map ψ

from (Y_2, d_2) to (X, d) such that $\phi = \psi \circ \pi$, which by part 1. is a non-expansive map from (Y_2, diam_{d_2}) to (X, diam_d) . Since $\delta_2(A) \geq \text{diam}_{d_2}(A)$ for all A , ψ is non-expansive from (Y_2, δ_2) to (X, diam_d) . Hence (X, diam_d) is injective.

Conversely, suppose (X, diam_d) is an injective diversity. Let $(Y_1, d_1), (Y_2, d_2)$ be two metric spaces, let π be an embedding of (Y_1, d_1) into (Y_2, d_2) , and let ϕ be a non-expansive map from (Y_1, d_1) to (X, d) . Then ϕ is a non-expansive map from (Y_1, diam_{d_1}) to (X, diam_d) and since (X, diam_d) is injective, there is a non-expansive map ψ from (Y_2, diam_{d_2}) to (X, diam_d) such that $\phi = \psi \circ \pi$. Applying part 1. again, we have that ψ is the required non-expansive map from (Y_2, d_2) to (X, d) . Hence (X, d) is injective.

3. Since (X, δ) is a diameter diversity, for any finite F and $\{A_f\}_{f \in F} \subseteq \mathcal{P}_{\text{fin}}(X)$, we have

$$\delta \left(\bigcup_{f \in F} A_f \right) = \delta(A_{f_1} \cup A_{f_2})$$

for some $f_1, f_2 \in F$. Hence for finite $F \subseteq T_X^\delta$

$$\begin{aligned} \delta_T(F) &= \sup_{A_f} \left\{ \delta \left(\bigcup_{f \in F} A_f \right) - \sum_{f \in F} f(A_f) \right\} \\ &= \max_{f_1, f_2 \in F} \sup_{A_1, A_2 \in \mathcal{P}_{\text{fin}}(X)} \{ \delta(A_1 \cup A_2) - f_1(A_1) - f_2(A_2) \} \\ &= \max_{f_1, f_2 \in F} \delta_T(\{f_1, f_2\}). \end{aligned}$$

□

Theorem 4.3. *Let (X, d) be a metric space with metric tight span (T_X^d, d_T) . Let (X, δ) be the associated diameter diversity where $\delta = \text{diam}_d$, and let (T_X^δ, δ_T) be its diversity tight span. Then*

1. *The metric space obtained by restricting δ_T to pairs in T_X^δ is isometric to the metric space (T_X^d, d_T) .*
2. *The diversity obtained by taking the diameter on the metric space (T_X^d, d_T) is isomorphic to the diversity (T_X, δ_T) .*

Proof.

First note that for any metric spaces (X_1, d_1) and (X_2, d_2) a map ϕ from X_1 to X_2 is an embedding from (X_1, d_1) to (X_2, d_2) if and only if ϕ is an embedding from (X_1, diam_{d_1}) to (X_2, diam_{d_2}) .

Let (T_X^d, δ_{d_T}) be the diameter diversity associated to (T_X^d, d_T) and let $(T_X^\delta, d_{\delta_T})$ be the induced metric for (T_X^δ, δ_T) . Let κ_d be the Kuratowski embedding from (X, d) to (T_X^d, d_T) . Then κ_d is also an embedding from (X, δ) to (T_X^d, δ_{d_T}) . In the same way, let κ_δ be the Kuratowski embedding from (X, δ) to (T_X^δ, δ_T) . Then κ_δ is also an embedding from (X, d) to $(T_X^\delta, d_{\delta_T})$.

By Lemma 4.2 2., (T_X^d, δ_{d_T}) is a hyperconvex diversity and $(T_X^\delta, d_{\delta_T})$ is a hyperconvex metric space. Applying [33, Proposition 9.20(4)] in the category **Met** there is an embedding ϕ from (T_X^d, d_T) to $(T_X^\delta, d_{\delta_T})$ such that

$$\kappa_\delta = \phi \circ \kappa_d. \quad (4.1)$$

The identity map $\text{id}_{T_X^d}$ on (T_X^d, d_T) is non-expansive and ϕ is an embedding, so applying the definition of injective metric spaces to (T_X^d, d_T) we have that there is a non-expansive map ψ from T_X^δ to T_X^d such that

$$\psi \circ \phi = \text{id}_{T_X^d}. \quad (4.2)$$

By Lemma 4.2 3., the diversity (T_X^δ, δ_T) is a diameter diversity and so from Lemma 4.2 1., the map ψ is also a non-expansive map from (T_X^δ, δ_T) to (T_X^d, δ_{d_T}) . Combining (4.1) and (4.2) we have

$$\begin{aligned} \psi \circ \kappa_\delta &= \psi \circ \phi \circ \kappa_d \\ &= \text{id}_{T_X^d} \circ \kappa_d \end{aligned}$$

which is an embedding. By Lemma 3.7 we have that ψ is an embedding, implying that ϕ is both an isomorphism from (T_X^d, d_T) to $(T_X^\delta, d_{\delta_T})$ and an isomorphism from (T_X^d, δ_{d_T}) to (T_X^δ, δ_T) . \square

5. Phylogenetic diversity

A metric space (X, d) is *additive* or *tree-like* if there is a tree with nodes partially labelled by X so that for each $x, y \in X$ the length of the path (including branch-lengths) connecting x and y equals $d(x, y)$. Dress [4] showed that if (X, d) is additive then its metric tight span corresponds exactly to the smallest tree it can be embedded in. The elements of the tight span correspond not only to the nodes of the original tree, but also the points along the edges. Here we will prove analogous results about phylogenetic diversity.

Following [4] we will work with *real trees* (also called metric-trees or \mathfrak{R} -trees), rather than graph-theoretic trees.

Definition 5.1. [34, 35]

1. Let (\mathcal{X}, d) be a metric space and let x, y be two points at distance $d(x, y) = r$. A geodesic joining x, y is a map $c : [0, r] \rightarrow \mathcal{X}$ such that $c(0) = x$, $c(r) = y$ and $d(c(s), c(t)) = |t - s|$ for all $s, t \in [0, r]$. The image of c is called a geodesic segment.
2. [34, Defn 2.1] A metric space (\mathcal{X}, d) is a real tree or \mathfrak{R} -tree if
 - (a) there is a unique geodesic segment $[x, y]$ joining each pair of points $x, y \in \mathcal{X}$.
 - (b) if $[y, x] \cap [x, z] = \{x\}$ then $[y, x] \cup [x, z] = [y, z]$.

Hence if x, y, z are three points in a real tree then

$$[x, y] \subseteq [x, z] \cup [y, z]. \quad (5.1)$$

Phylogenetic diversity, as introduced by [23] and investigated extensively by [8, 24, 25] and others, can be viewed as a generalisation of additive metrics. The phylogenetic diversity of a set of nodes or points in a tree is the length of the smallest subtree connecting them, so that the restriction of a phylogenetic diversity to pairs of points gives an additive metric. A formal definition of phylogenetic diversity on real trees requires a bit more machinery.

For a real tree (\mathcal{X}, d) , let μ be the one-dimensional Hausdorff measure on it [36]. The important features of μ for our purposes is that it is defined on all

Borel sets, it is monotone, and it is additive on disjoint sets. Furthermore, for any points $a, b \in \mathcal{X}$, $\mu([a, b]) = d(a, b)$, and naturally $\mu(\{a\}) = 0$. See [37] for a related measure on real trees.

Definition 5.2. 1. *The convex hull of a set $A \subseteq \mathcal{X}$ is*

$$\text{conv}(A) = \bigcup_{a, b \in A} [a, b]$$

and we say that A is convex if $A = \text{conv}(A)$.

2. *Let (\mathcal{X}, d) be a real tree. The real-tree diversity (\mathcal{X}, δ_t) for (\mathcal{X}, d) is defined by*

$$\delta_t(A) := \mu(\text{conv}(A))$$

for all finite $A \subseteq \mathcal{X}$. Note that since A is finite, $\text{conv}(A)$ is closed and hence $\mu(\text{conv}(A))$ is defined.

First we prove that this phylogenetic diversity satisfies the diversity axioms (D1) and (D2).

Theorem 5.3. *Let (\mathcal{X}, d) be a real tree. Then (\mathcal{X}, δ_t) is a diversity.*

Proof.

Since μ is a measure, δ_t is non-negative and also monotonic. If $|A| \leq 1$ then $\text{conv}(A) = A$ and so $\delta_t(A) = \mu(A) = 0$. If $|A| > 1$ then select distinct $a, b \in A$. Since $\text{conv}([a, b]) = [a, b]$ and $\mu([a, b]) = d(a, b)$ we have $\delta_t(A) \geq \delta_t(\{a, b\}) = d(a, b) > 0$. This proves (D1).

Let $A, B, C \in \mathcal{P}_{\text{fin}}(\mathcal{X})$ and suppose that $B \neq \emptyset$. From (5.1) we have

$$[a, c] \subseteq [a, b] \cup [b, c] \tag{5.2}$$

for all $a \in A$, $b \in B$ and $c \in C$. Hence

$$\text{conv}(A \cup C) \subseteq \text{conv}(A \cup B) \cup \text{conv}(B \cup C)$$

and

$$\begin{aligned} \delta_t(A \cup C) &= \mu(\text{conv}(A \cup C)) \\ &\leq \mu(\text{conv}(A \cup B)) + \mu(\text{conv}(B \cup C)) \\ &= \delta_t(A \cup B) + \delta_t(B \cup C), \end{aligned}$$

giving us the triangle equality (D2). □

We now show that complete real-tree diversities are hyperconvex, proving the diversity analogue of [38, Theorem 3.2].

Lemma 5.4. *Let (\mathcal{X}, d) be a real tree with associated tree diversity (\mathcal{X}, δ_t) . For all finite $C \subseteq \mathcal{X}$ and $r \geq \delta_t(C)$, the ball $B(C, r) = \{x \in \mathcal{X} : \delta_t(C \cup \{x\}) \leq r\}$ is closed and convex.*

Proof.

For any finite but non-empty $C \subseteq \mathcal{X}$ the function

$$\phi : \mathcal{X} \rightarrow \mathbb{R} : x \mapsto \delta_t(C \cup \{x\})$$

is continuous. Hence when $r \geq \delta_t(C)$ the ball

$$B(C, r) := \phi^{-1}(A) = \{x \in T_X^d : \delta_t(C \cup \{x\}) \leq r\}$$

is closed.

To prove convexity, suppose that $x_1, x_2 \in B(C, r)$. Fix $a \in C$. For all $y \in [a, x_1]$, $\text{conv}(C \cup \{y\}) \subseteq \text{conv}(C \cup \{x_1\})$ and so $\delta_t(C \cup \{y\}) \leq \delta_t(C \cup \{x_1\})$ showing that $y \in B(C, r)$. We have that $[a, x_1]$, and by symmetry $[a, x_2]$, are contained in $B(C, r)$. By (5.1) we have

$$[x_1, x_2] \subseteq [a, x_1] \cup [a, x_2] \subseteq B(C, r)$$

so that $B(C, r)$ is both closed and convex. □

Theorem 5.5. *Let (\mathcal{X}, d) be a real tree with associated real-tree diversity (\mathcal{X}, δ_t) . Then (\mathcal{X}, δ_t) is hyperconvex if and only if (\mathcal{X}, d) is complete.*

Proof.

Suppose that (\mathcal{X}, d) is a complete real tree. Then (\mathcal{X}, d) is a hyperconvex metric space [38, Theorem 3.2]. Suppose that $r : \mathcal{P}_{\text{fin}}(\mathcal{X}) \rightarrow \mathfrak{R}$ satisfies

$$\delta_t \left(\bigcup_{A \in \mathcal{A}} A \right) \leq \sum_{A \in \mathcal{A}} r(A)$$

for all finite $\mathcal{A} \subseteq \mathcal{P}_{\text{fin}}(\mathcal{X})$. We will show that the collection of balls

$$\Gamma = \{B(A, r(A)) : A \in \mathcal{P}_{\text{fin}}(\mathcal{X})\}$$

has a non-empty intersection.

First we show that the members of Γ intersect pairwise. Consider a pair of nonempty finite subsets A_i, A_j of X . To show that $B(A_i, r(A_i))$ and $B(A_j, r(A_j))$ intersect, we show that there is v such that $\delta_t(A_i \cup \{v\}) \leq r(A_i)$ and $\delta_t(A_j \cup \{v\}) \leq r(A_j)$. This clearly holds if there is $\text{conv}(A_i) \cap \text{conv}(A_j) \neq \emptyset$. Suppose then that $\text{conv}(A_i)$ and $\text{conv}(A_j)$ are disjoint. Since A_i, A_j are finite, $\text{conv}(A_i)$ and $\text{conv}(A_j)$ are closed subtrees of T_X^d . By [35, Ch. 2, Lemma 1.9] there exists $a_i \in \text{conv}(A_i)$ and $a_j \in \text{conv}(A_j)$ such that $[a_i, a_j] \cap \text{conv}(A_i) = \{a_i\}$ and $[a_i, a_j] \cap \text{conv}(A_j) = \{a_j\}$ and for all $x \in A_i$ and $y \in A_j$ we have $[a_i, a_j] \subseteq [x, y]$. Then,

$$\begin{aligned} r(A_i) + r(A_j) &\geq \delta_t(A_i \cup A_j) \\ &= \mu(\text{conv}(A_i \cup A_j)) \\ &\geq \mu(\text{conv}(A_i)) + \mu([a_i, a_j]) + \mu(\text{conv}(A_j)) \\ &= \delta_t(A_i) + d(a_i, a_j) + \delta_t(A_j). \end{aligned}$$

Hence there is $v \in [a_i, a_j]$ such that $d(a_i, v) \leq r(A_i) - \delta_t(A_i)$ and $d(a_j, v) \leq r(A_j) - \delta_t(A_j)$, so that

$$\begin{aligned} \delta_t(A_i \cup \{v\}) &= \delta_t(A_i) + \delta_t(\{a_i, v\}) \\ &= \delta_t(A_i) + d(a_i, v) \\ &\leq r(A_i), \end{aligned}$$

and likewise $\delta_t(A_j \cup \{v\}) \leq r(A_j)$.

We have established that Γ satisfies the pairwise intersection property. The closed, convex sets of a real tree satisfy the Helly property [34], so every finite subcollection of Γ has non-empty intersection. By the completeness of (\mathcal{X}, d) , Γ has a non-empty intersection, so there is v such that $\delta_t(A \cup \{v\}) \leq r(A)$ for all $A \in \mathcal{P}_{\text{fin}}(\mathcal{X})$. This proves that (\mathcal{X}, δ_t) is hyperconvex.

For the converse, we note that completeness of (\mathcal{X}, d) follows directly from [2, Proposition 3.2] and the definition of hyperconvexity for diversities. \square

Definition 5.6. *A diversity (X, δ) is a phylogenetic diversity if it can be embedded in a real-tree diversity (\mathcal{X}, δ_t) for some complete real tree (\mathcal{X}, d) .*

Clearly, every real-tree diversity is a phylogenetic diversity, but a phylogenetic diversity is a real-tree diversity only if its induced metric is a real tree.

Theorem 5.7. *Let (X, δ) be a diversity. Then (X, δ) is a phylogenetic diversity if and only if (T_X, δ_T) is a real-tree diversity.*

Proof.

Since (X, δ) is a phylogenetic diversity there is a complete real tree (\mathcal{X}, d) with real-tree diversity (\mathcal{X}, δ_t) for which there is an embedding ϕ from (X, δ) into (\mathcal{X}, δ_t) . By Theorem 5.5 (\mathcal{X}, δ_t) is hyperconvex. By Theorem 3.8 there is an embedding ψ from (T_X, δ_T) into (\mathcal{X}, δ_t) such that $\phi = \psi \circ \kappa$.

Let (T_X, d_{δ_T}) be the induced metric for (T_X, δ_T) . It follows directly from the hyperconvexity of (T_X, δ_T) that (T_X, d_{δ_T}) is convex. For any $f, g \in T_X$ and geodesic segment $[f, g]$ in T_X , the image of $[f, g]$ under ψ is the unique geodesic segment between $\psi(f)$ and $\psi(g)$. It follows that $\psi(T_X)$ is a convex subset of (\mathcal{X}, d) and (\mathcal{X}, d) restricted to $\psi(T_X)$ is a real tree [35, pg. 36]. Restricting (X, δ_t) to $\psi(T_X)$ then gives a real-tree diversity which is isomorphic to (T_X, δ_T) .

For the converse, note that the map κ from (X, δ) into its tight span is an embedding, so that (X, δ) is a phylogenetic diversity. \square

We now link the real tree given by the diversity tight span of a phylogenetic diversity and the tight span of its induced metric.

Lemma 5.8. *Let (X, d_X) and (Y, d_Y) be complete real trees and let (X, δ_X) and (Y, δ_Y) be the associated real-tree diversities. Then*

1. *$\psi : X \rightarrow Y$ is a non-expansive map from (X, d_X) to (Y, d_Y) if and only if it is a non-expansive map from (X, δ_X) to (Y, δ_Y) .*
2. *$\psi : X \rightarrow Y$ is an embedding from (X, d_X) to (Y, d_Y) if and only if it is an embedding from (X, δ_X) to (Y, δ_Y) .*

Proof.

1. Suppose that $\psi : X \rightarrow Y$ is a non-expansive map from (X, d_X) to (Y, d_Y) . For any finite $A \subset X$, we have $\delta_Y(\psi(A)) = \mu(\text{conv}(\psi(A)))$. First note that $\text{conv}(\psi(A)) = \psi(\text{conv}(A))$. Then note that since in this case the one-dimensional Hausdorff measure of a set is a limit of infima of the total length of countable covers of a set by geodesic segments [36, Section 6.1], $\mu(\psi(B)) \leq \mu(B)$ for all measurable $B \subseteq X$. This proves that ψ is a non-expanding map with respect to diversities. The other direction is immediate.
2. The argument follows as in 1, with showing $\mu(\psi(B)) = \mu(B)$ for all measurable $B \subseteq X$. □

Theorem 5.9. *Let (X, δ) be a phylogenetic diversity and let (X, d) be its induced metric. Let (T_X^δ, δ_T) be the diversity tight span of (X, δ) and let (T_X^d, d_T) be the metric tight span of (X, d) . Then (T_X^d, d_T) is isometric with the induced metric of (T_X^δ, δ_T) .*

Proof.

By [4, Theorem 8], (T_X^d, d_T) is a real tree. Let (T_X^d, δ_{d_T}) be the corresponding real-tree diversity, which is hyperconvex by Theorem 5.5. Let $(T_X^\delta, \delta_\delta)$ denote the induced metric of (T_X^δ, δ_T) . From Theorem 5.5 we have that $(T_X^\delta, \delta_\delta)$ is a complete real tree and is therefore a hyperconvex metric space [38, Theorem 3.2].

Let κ_d be the Kuratowski embedding from (X, d) to (T_X^d, d_T) and let κ_δ be the Kuratowski embedding from (X, δ) to (T_X^δ, δ_T) . The map κ_δ is then also an

embedding between the induced metric (X, d) and the induced metric $(T_X^\delta, d_{\delta_T})$. Applying [33, Proposition 9.20(4)] in the category **Met**, there is an embedding $\phi : (T_X^d, d_T) \rightarrow (T_X^\delta, d_{\delta_T})$ such that

$$\kappa_\delta = \phi \circ \kappa_d. \quad (5.3)$$

By Lemma 5.8, ϕ is also an embedding from the diversity (T_X^d, δ_{d_T}) to the diversity (T_X^δ, δ_T) . For all $A \in \mathcal{P}_{\text{fin}}(X)$,

$$\delta(A) = \delta_T(\kappa_\delta(A)) = \delta_T(\phi(\kappa_d(A))) = \delta_{d_T}(\kappa_d(A))$$

so that κ_d embeds (X, δ) in (T_X^d, δ_{d_T}) .

The identity map $\text{id}_{T_X^d}$ on (T_X^d, d_T) is non-expansive and ϕ is an embedding, so applying the definition of injective metric spaces to (T_X^d, d_T) we have that there is a non-expansive map ψ from $(T_X^\delta, d_{\delta_T})$ to (T_X^d, d_T) such that

$$\psi \circ \phi = \text{id}_{T_X^d}. \quad (5.4)$$

Applying Lemma 5.8 1. we see that the map ψ is also a non-expansive map from (T_X^δ, δ_T) to (T_X^d, δ_{d_T}) . Combining (5.3) and (5.4) we have

$$\begin{aligned} \psi \circ \kappa_\delta &= \psi \circ \phi \circ \kappa_d \\ &= \text{id}_{T_X^d} \circ \kappa_d \end{aligned}$$

which is an embedding. By Lemma 3.7 we have that ψ is an embedding, implying that ϕ is an isometry from (T_X^d, d_T) to $(T_X^\delta, d_{\delta_T})$. \square

6. Tight span and the Steiner tree problem

Let X be a finite set of points in a metric space (M, d) . The *(metric) Steiner tree problem* is to find the shortest network that connects them. Clearly this network will always be a tree. More formally

METRIC STEINER PROBLEM.

Input: Subset X of a metric space (M, d) .

Problem: Find a (graph theoretic) tree T for which $X \subseteq V(T) \subseteq M$ and

$$\sum_{\{u,v\} \in E(T)} d(u,v)$$

is minimised.

Dress and Krüger [32] examined an ‘abstract’ metric Steiner problem where one drops the constraint that $V(T) \subseteq M$. This abstract Steiner tree was one of the first distance-based criteria proposed for the inference of phylogenetic trees [39, 40], though it is now not widely used. Suppose that T is a tree with edge weights $w : E(T) \rightarrow \mathbb{R}_{\geq 0}$. Given $u, v \in V(T)$ we let $d_w(u, v)$ denote the sum of edge weights along the path from u to v .

ABSTRACT STEINER PROBLEM.

Input: Finite metric space (X, d) .

Problem: Find a (graph theoretic) tree T and edge weighting $w : E(T) \rightarrow \mathbb{R}$ such that $X \subseteq V(T)$, $d_w(x, y) \geq d(x, y)$ for all $x, y \in X$ and

$$\sum_{e \in E(T)} w(e)$$

is minimised.

Suppose that T is a solution to the metric Steiner problem for $X \subseteq M$. Define the weight function $w : E(T) \rightarrow \mathbb{R}$ by $w(\{u, v\}) = d(u, v)$. Then, by the triangle inequality, $d_w(x, y) \geq d(x, y)$ for all $x, y \in X$. It follows then that the length of the minimum abstract Steiner tree for $(X, d|_X)$ is a lower bound for the metric Steiner problem. Dress and Krüger showed that the lower bound becomes tight when (M, d) equals (T_X, d_T) , the metric tight span of X .

Theorem 6.1 ([32]). *Let (X, d) be a finite metric space. For every solution (T, w) to the abstract Steiner tree problem there is a map $\phi : V(T) \rightarrow T_X$ such that $\phi(x) = \kappa(x)$ for all $x \in X$ and $w(\{u, v\}) = d_T(\phi(u), \phi(v))$ for all $\{u, v\} \in E(T)$.*

Hence the length of the minimal Steiner tree for $\kappa(X)$ in (T_X, d_T) equals the length of the minimal abstract Steiner tree for (X, d) and the minimal abstract Steiner trees can be embedded within the tight span. A direct corollary is that if d is tree-like then the abstract Steiner tree equals the tree corresponding to d .

Here we show that, using diversities, we can obtain a tighter bound on the metric Steiner problem than that given by the abstract Steiner problem. Given a tree T with edge weights w and $A \subseteq V(T)$ we let $\delta_w(A)$ be the sum of edge weights in the smallest subtree of T connecting A . Hence $(X, \delta_w|_X)$ is a phylogenetic diversity.

DIVERSITY STEINER PROBLEM.

Input: Finite diversity (X, δ) .

Problem: Find a (graph theoretic) tree T and edge weighting $w : E(T) \rightarrow \mathbb{R}$ such that $X \subseteq V(T)$, $\delta_w(Y) \geq \delta(Y)$ for all $Y \subseteq X$, and

$$\sum_{e \in E(T)} w(e)$$

is minimised.

Let X be a finite subset of a metric space (M, d) . For each $A \subseteq X$ let $\ell(A)$ denote the minimum length of a (metric) Steiner tree connecting the points A in the metric space (M, d) . We see that (X, ℓ) is a diversity. For each $k \geq 2$, consider the *truncated* diversity $\delta^{(k)}$ defined by

$$\delta^{(k)}(A) = \max\{\ell(B) : |B| \leq k, B \subseteq A\}$$

for all $A \subseteq X$.

Proposition 6.2. *If (T, w) is a minimum length solution for the diversity Steiner problem applied to $\delta^{(k)}$ then the length $\sum_{e \in E(T)} w(e)$ of T is a lower bound for $\ell(X)$, the optimal length of a metric Steiner tree for X .*

Proof.

Let (T', w') be a solution to the metric Steiner problem and let $\delta_{w'}$ be the

associated phylogenetic diversity. Then for all B such that $|B| \leq k$ we have that $\delta_{w'}(B)$, the length of T' restricted to B , is bounded below by $\ell(B) = \delta^{(k)}(B)$. It follows that $\delta^{(k)}(A) \leq \delta_{w'}(A)$ for all $A \subseteq X$, so that (T', w') is a potential solution for the diversity Steiner problem. As (T, w) is optimal, we have

$$\sum_{e \in E(T)} w(e) \leq \sum_{e \in E(T')} w'(e) = \ell(X).$$

□

For $k = 2$, the bounds provided by the Proposition 6.2 coincide with those given by length of the minimum abstract Steiner tree. As k increases, the bounds returned by the diversity Steiner tree applied to $\delta^{(k)}$ will tighten, until eventually the diversity Steiner tree will coincide with the metric Steiner tree. Furthermore, we have a direct extension of Theorem 6.1, stating that these diversity Steiner trees will all be contained in the diversity tight span.

Theorem 6.3. *Let (X, δ) be a finite diversity. For every solution (T, w) to the diversity Steiner tree problem for (X, δ) there is a map $\phi : V(T) \rightarrow T_X$ such that $\phi(x) = \kappa(x)$ for all $x \in X$ and $w(\{u, v\}) = \delta_T(\{\phi(u), \phi(v)\})$ for all $\{u, v\} \in E(T)$.*

Proof.

Let δ_w be the diversity on $V(T)$ given by (T, w) , as defined above. Since (T, w) solves the diversity Steiner problem, $\delta_w(A) \geq \delta(A)$ for all $A \subseteq X$. Let κ denote the canonical embedding from X to T_X . Then κ is a non-expansive map from $(X, \delta_w|_X)$ to (T_X, δ_T) .

The tight span (T_X, δ_T) is injective. Hence there is a non-expansive map ϕ from $(V(T), \delta_w)$ to (T_X, δ_T) such that $\phi(x) = \kappa(x)$ for all $x \in X$. For each u, v let $w'(\{u, v\}) = \delta_T(\{\phi(u), \phi(v)\})$. Then

$$w(\{u, v\}) = \delta_w(\{u, v\}) \geq \delta_T(\{\phi(u), \phi(v)\}) = w'(\{u, v\})$$

for all $u, v \in V$.

Consider $A \subseteq X$, and let E_A be the set of edges in the smallest subtree of T containing A . By the triangle inequality,

$$\delta_{w'}(A) = \sum_{e \in E_A} w'(e) \geq \delta(X).$$

Hence (T, w') is a candidate for the diversity Steiner problem, but since (T, w) is already minimum, $\sum_{e \in E(T)} w(e) \leq \sum_{e \in E(T)} w'(e)$. It follows that $w(e) = w'(e)$ for all $e \in E(T)$. \square

References

- [1] N. Aronszajn, P. Panitchpakdi, Extension of uniformly continuous transformations and hyperconvex metric spaces, *Pacific J. Math.* 6 (1956) 405–439.
- [2] R. Espínola, M. A. Khamisi, Introduction to hyperconvex spaces, in: *Handbook of metric fixed point theory*, Kluwer Acad. Publ., Dordrecht, 2001, pp. 391–435.
- [3] J. R. Isbell, Six theorems about injective metric spaces, *Comment. Math. Helv.* 39 (1964) 65–76.
- [4] A. W. M. Dress, Trees, tight extensions of metric spaces, and the cohomological dimension of certain groups: a note on combinatorial properties of metric spaces, *Adv. Math.* 53 (3) (1984) 321–402.
- [5] H.-J. Bandelt, A. W. M. Dress, A canonical decomposition theory for metrics on a finite set, *Adv. Math.* 92 (1) (1992) 47–105.
- [6] D. Bryant, V. Moulton, NeighborNet: An agglomerative algorithm for the construction of planar phylogenetic networks, *Mol. Biol. Evol.* 21 (2004) 255–265.
- [7] D. Huson, D. Bryant, Application of phylogenetic networks in evolutionary studies, *Mol. Biol. Evol.* 23 (2006) 254–267.

- [8] L. Pachter, D. Speyer, Reconstructing trees from subtree weights, *Appl. Math. Lett.* 17 (6) (2004) 615–621.
- [9] A. Dress, V. Moulton, W. Terhalle, *T*-theory: an overview, *European J. Combin.* 17 (2-3) (1996) 161–175.
- [10] B. Sturmfels, Can biology lead to new theorems?, *Annual report of the Clay Mathematics Institute* (2005) 13–26.
- [11] A. Dress, K. T. Huber, A. Lesser, V. Moulton, Hereditarily optimal realizations of consistent metrics, *Ann. Comb.* 10 (1) (2006) 63–76.
- [12] K. T. Huber, J. Koolen, V. Moulton, A. Spillner, Characterizing cell-decomposable metrics, *Electron. J. Combin.* 15 (1) (2008) Note 7, 9.
- [13] P. Buneman, A note on the metric properties of trees, *J. Combinatorial Theory Ser. B* 17 (1974) 48–50.
- [14] A. Dress, K. Huber, V. Moulton, Some variations on a theme by Buneman, *Ann. Comb.* 1 (4) (1997) 339–352.
- [15] A. Dress, M. Hendy, K. Huber, V. Moulton, On the number of vertices and edges of the Buneman graph, *Ann. Comb.* 1 (4) (1997) 329–337.
- [16] K. T. Huber, J. H. Koolen, V. Moulton, On the structure of the tight-span of a totally split-decomposable metric, *European J. Combin.* 27 (3) (2006) 461–479.
- [17] M. Develin, B. Sturmfels, Tropical convexity, *Doc. Math.* 9 (2004) 1–27 (electronic).
- [18] P. Huggins, B. Sturmfels, J. Yu, D. S. Yuster, The hyperdeterminant and triangulations of the 4-cube, *Math. Comp.* 77 (263) (2008) 1653–1679.
- [19] B. Sturmfels, J. Yu, Classification of six-point metrics, *Electron. J. Combin.* 11 (1) (2004) Research Paper 44, 16 pp. (electronic).

- [20] K. T. Huber, J. H. Koolen, V. Moulton, The tight span of an antipodal metric space. I. Combinatorial properties, *Discrete Math.* 303 (1-3) (2005) 65–79.
- [21] D. Eppstein, Manhattan orbifolds, *Topology Appl.* 157 (2) (2010) 494–507.
- [22] H. Hirai, A geometric study of the split decomposition, *Discrete Comput. Geom.* 36 (2) (2006) 331–361.
- [23] D. Faith, Conservation evaluation and phylogenetic diversity, *Biological Conservation* 61 (1992) 1–10.
- [24] M. A. Steel, Phylogenetic diversity and the greedy algorithm, *Syst. Biol.* 54 (4) (2005) 527–529.
- [25] B. Minh, S. Klaere, A. von Haeseler, Taxon selection under split diversity, *Syst. Biol.* 58 (6) (2009) 586–594.
- [26] S. Gähler, Untersuchungen über verallgemeinerte m -metrische Räume, I, II, III, *Math. Nachr.* 40 (1969), 165–189; *ibid.* 40 (1969), 229–264; *ibid.* 41 (1969) 23–36.
- [27] K. Menger, Untersuchungen über allgemeine Metrik, *Math. Ann.* 100 (1) (1928) 75–163.
- [28] M.-M. Deza, I. G. Rosenberg, Small cones of m -hemimetrics, *Discrete Math.* 291 (1-3) (2005) 81–97.
- [29] V. Chepoi, B. Fichet, A note on three-way dissimilarities and their relationship with two-way dissimilarities, in: *Selected contributions in data analysis and classification*, *Stud. Classification Data Anal. Knowledge Organ.*, Springer, Berlin, 2007, pp. 465–475.
- [30] M. J. Warrens, n -way metrics, *J. Classification* 27 (2010) In press.
- [31] A. Dress, W. Terhalle, The tree of life and other affine buildings, in: *Proceedings of the International Congress of Mathematicians*, Vol. III (Berlin, 1998), 1998, pp. 565–574.

- [32] A. Dress, M. Krüger, Parsimonious phylogenetic trees in metric spaces and simulated annealing, *Adv. in Appl. Math.* 8 (1) (1987) 8–37.
- [33] J. Adámek, H. Herrlich, G. E. Strecker, Abstract and concrete categories, Pure and Applied Mathematics (New York), John Wiley & Sons Inc., New York, 1990, the joy of cats, A Wiley-Interscience Publication.
- [34] R. Espínola, W. A. Kirk, Fixed point theorems in \mathbb{R} -trees with applications to graph theory, *Topology Appl.* 153 (7) (2006) 1046–1055.
- [35] I. Chiswell, Introduction to Λ -trees, World Scientific Publishing Co. Inc., River Edge, NJ, 2001.
- [36] G. Edgar, Measure, topology, and fractal geometry, 2nd Edition, Undergraduate Texts in Mathematics, Springer, New York, 2008.
- [37] S. N. Evans, A. Winter, Subtree prune and regraft: a reversible real tree-valued Markov process, *Ann. Probab.* 34 (3) (2006) 918–961.
- [38] W. A. Kirk, Hyperconvexity of \mathbf{R} -trees, *Fund. Math.* 156 (1) (1998) 67–72.
- [39] W. Beyer, M. Stein, T. Smith, S. Ulam, A molecular sequence metric and evolutionary trees, *Math. Biosci.* 19 (1974) 9–25.
- [40] M. Waterman, T. Smith, M. Singh, W. Beyer, Additive evolutionary trees, *J. Theor. Biol.* 64 (1977) 199–213.